A Tseng's type penalty scheme for solving inclusion problems involving linearly composed and parallel-sum type monotone operators

Radu Ioan Boţ [*] Ernö Robert Csetnek [†]

September 25, 2013

Dedicated to Professor Boris Mordukhovich on the occasion of his 65th birthday

Abstract. In this paper we consider the inclusion problem involving a maximally monotone operator, a monotone and Lipschitz continuous operator, linear compositions of parallel-sum type monotone operators as well as the normal cone to the set of zeros of another monotone and Lipschitz continuous operator. We propose a forward-backward-forward type algorithm for solving it that assumes an individual evaluation of each operator. Weak ergodic convergence of the sequence of iterates generated by the algorithmic scheme is guaranteed under a condition formulated in terms of the Fitzpatrick function associated to one of the monotone and Lipschitz continuous operators. We also discuss show how the proposed penalty scheme can be applied to convex minimization problems and present some numerical experiments in TV-based image inpainting.

Key Words. maximally monotone operator, Fitzpatrick function, resolvent, Lipschitz continuous operator, parallel-sum, forward-backward-forward algorithm, subdifferential, Fenchel conjugate, infimal-convolution, convex minimization problem

AMS subject classification. 47H05, 65K05, 90C25

1 Introduction and preliminaries

1.1 Motivation

In the last couple of years a number of papers have been published dealing with the solving of monotone inclusion problems of the form

$$0 \in Ax + N_M(x), \tag{1}$$

[*]University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: radu.bot@univie.ac.at. Research partially supported by DFG (German Research Foundation), project BO 2516/4-1.

[†]University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, e-mail: ernoe.robert.csetnek@univie.ac.at. Research supported by DFG (German Research Foundation), project BO 2516/4-1.

where $\mathcal{H}$ is a real Hilbert space, $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximally monotone operator, $M = \operatorname{argmin} \Psi$ is the set of global minima of the proper, convex and lower semicontinuous function $\Psi : \mathcal{H} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ fulfilling $\min \Psi = 0$ and $N_M : \mathcal{H} \rightrightarrows \mathcal{H}$ is the normal cone to the set $M \subseteq \mathcal{H}$ (see [1–3, 17, 18]).

When A is the convex subdifferential of a proper, convex and lower semicontinuous function $\Phi : \mathcal{H} \to \overline{\mathbb{R}}$ fulfilling an appropriate qualification condition, then this gives rise to the solving of a convex minimization problem of the form

$$\min_{x \in \mathcal{H}} \{\Phi(x) : x \in \operatorname{argmin} \Psi\}. \tag{2}$$

The algorithms given in the mentioned literature in the context of solving (1) are forward-backward type penalty schemes, performing in each iteration a proximal step with respect to A and a subgradient step with respect to the penalization function Ψ. Convergence results were usually proven assuming that

$$\text{for every } p \in \operatorname{ran} N_M, \sum_{n \in \mathbb{N}} \lambda_n \beta_n \left[\Psi^* \left(\frac{p}{\beta_n} \right) - \sigma_M \left(\frac{p}{\beta_n} \right) \right] < +\infty, \tag{3}$$

which is basically the discrete version of a condition given in the continuous case for nonautonomous differential inclusions in [1]. Here, $\Psi^* : \mathcal{H} \to \overline{\mathbb{R}}$ denotes the Fenchel conjugate function of Ψ, $\operatorname{ran} N_M$ the range of the normal cone operator $N_M : \mathcal{H} \rightrightarrows \mathcal{H}$, σ_M the support function of M and $(\lambda_n)_{n \in \mathbb{N}}$ and $(\beta_n)_{n \in \mathbb{N}}$ are positive real sequences that appear in the algorithm. For conditions guaranteeing (3) we refer the reader to [1–3,17,18].

In [10] we investigated the more general inclusion problem

$$0 \in Ax + Dx + N_M(x), \tag{4}$$

where $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximally monotone operator, $D : \mathcal{H} \to \mathcal{H}$ a single-valued cocoercive (respectively, monotone and Lipschitz continuous) operator and $M \subseteq \mathcal{H}$ the nonempty set of zeros of another cocoercive (respectively, monotone and Lipschitz continuous) operator $B : \mathcal{H} \to \mathcal{H}$. We formulated a forward-backward type and a Tseng's type algorithm for solving these problems and we also generalized (3) to a condition guaranteeing weak ergodic convergence for the sequence of generated iterates, which we formulated by using the Fitzpatrick function associated to B.

As a continuation of these developments, we deal in this paper with the solving of monotone inclusion problems having a more complex structure. We consider the problem of finding the zeros of a sum of maximally monotone operator with a monotone and Lipschitz continuous one, with the linear composition of parallel-sum type monotone operators and with the normal cone to the set of zeros of another monotone and Lipschitz operator. We propose a forward-backward-forward type penalty scheme for solving this inclusion problem. The proof of the convergence result relies on the fruitful idea that the inclusion problem under investigation can be written as a problem of type (4) in an appropriate product space. This will be basically done in the next section. In Section 3 we employ the outcomes of Section 2 in the context of solving convex minimization problems with intricate objective functions. Finally, in the last section of the paper we consider a numerical example in image inpainting.

1.2 Notations and preliminary results

For the readers convenience we present first some notations which are used throughout the paper (see [4, 7, 8, 15, 22, 23]). By $\mathbb{N} = \{1, 2, ...\}$ we denote the set of *positive integer numbers* and let $\mathcal{H}$ be a real Hilbert space with *inner product* $\langle \cdot, \cdot \rangle$ and associated *norm* $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$. When $\mathcal{G}$ is another Hilbert space and $L : \mathcal{H} \to \mathcal{G}$ a linear continuous operator, then the *norm* of L is defined as $\|L\| = \sup\{\|Lx\| : x \in \mathcal{H}, \|x\| \leq 1\}$, while $L^* : \mathcal{G} \to \mathcal{H}$, defined by $\langle L^*y, x \rangle = \langle y, Lx \rangle$ for all $(x, y) \in \mathcal{H} \times \mathcal{G}$, denotes the *adjoint operator* of L.

For a function $f : \mathcal{H} \to \overline{\mathbb{R}}$ we denote by $\operatorname{dom} f = \{x \in \mathcal{H} : f(x) < +\infty\}$ its *effective domain* and say that f is *proper*, if $\operatorname{dom} f \neq \emptyset$ and $f(x) \neq -\infty$ for all $x \in \mathcal{H}$. Let $f^* : \mathcal{H} \to \overline{\mathbb{R}}$, $f^*(u) = \sup_{x\in\mathcal{H}}\{\langle u, x \rangle - f(x)\}$ for all $u \in \mathcal{H}$, be the *conjugate function* of f. We denote by $\Gamma(\mathcal{H})$ the family of proper convex and lower semi-continuous extended real-valued functions defined on $\mathcal{H}$. The *subdifferential* of f at $x \in \mathcal{H}$, with $f(x) \in \mathbb{R}$, is the set $\partial f(x) := \{v \in \mathcal{H} : f(y) \geq f(x) + \langle v, y - x \rangle \ \forall y \in \mathcal{H}\}$. We take by convention $\partial f(x) := \emptyset$, if $f(x) \in \{\pm\infty\}$. We also denote by $\min f := \inf_{x\in\mathcal{H}} f(x)$ and by $\operatorname{argmin} f := \{x \in \mathcal{H} : f(x) = \min f\}$. For $f, g : \mathcal{H} \to \overline{\mathbb{R}}$ two proper functions, we consider their *infimal convolution*, which is the function $f \Box g : \mathcal{H} \to \overline{\mathbb{R}}$, defined by $(f \Box g)(x) = \inf_{y\in\mathcal{H}}\{f(y) + g(x - y)\}$, for all $x \in \mathcal{H}$.

Let $M \subseteq \mathcal{H}$ be a nonempty set. The *indicator function* of M, $\delta_M : \mathcal{H} \to \overline{\mathbb{R}}$, is the function which takes the value 0 on M and $+\infty$ otherwise. The subdifferential of the indicator function is the *normal cone* of M, that is $N_M(x) = \{u \in \mathcal{H} : \langle u, y - x \rangle \leq 0 \ \forall y \in M\}$, if $x \in M$ and $N_M(x) = \emptyset$ for $x \notin M$. Notice that for $x \in M$, $u \in N_M(x)$ if and only if $\sigma_M(u) = \langle u, x \rangle$, where σ_M is the support function of M, defined by $\sigma_M(u) = \sup_{y\in M}\langle y, u \rangle$. If $M \subseteq \mathcal{H}$ is a convex set, we denote by

$$\operatorname{sqri} M := \{x \in M : \cup_{\lambda>0}\lambda(M - x) \text{ is a closed linear subspace of } \mathcal{H}\}$$

its *strong quasi-relative interior*. Notice that we always have $\operatorname{int} M \subseteq \operatorname{sqri} M$ (in general this inclusion may be strict). If $\mathcal{H}$ is finite-dimensional, then $\operatorname{sqri} M$ coincides with $\operatorname{ri} M$, the relative interior of M, which is the interior of M with respect to its affine hull.

For an arbitrary set-valued operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ we denote by $\operatorname{Gr} A = \{(x, u) \in \mathcal{H} \times \mathcal{H} : u \in Ax\}$ its *graph*, by $\operatorname{dom} A = \{x \in \mathcal{H} : Ax \neq \emptyset\}$ its *domain*, by $\operatorname{ran} A = \{u \in \mathcal{H} : \exists x \in \mathcal{H} \text{ s.t. } u \in Ax\}$ its *range* and by $A^{-1} : \mathcal{H} \rightrightarrows \mathcal{H}$ its *inverse operator*, defined by $(u, x) \in \operatorname{Gr} A^{-1}$ if and only if $(x, u) \in \operatorname{Gr} A$. The *parallel sum* of two set-valued operators $A_1, A_2 : \mathcal{H} \rightrightarrows \mathcal{H}$ is defined as

$$A_1 \Box A_2 : \mathcal{H} \rightrightarrows \mathcal{H}, A_1 \Box A_2 = \left(A_1^{-1} + A_2^{-1}\right)^{-1}.$$

We use also the notation $\operatorname{zer} A = \{x \in \mathcal{H} : 0 \in Ax\}$ for the *set of zeros* of the operator A. We say that A is *monotone* if $\langle x - y, u - v \rangle \geq 0$ for all $(x, u), (y, v) \in \operatorname{Gr} A$. A monotone operator A is said to be *maximally monotone*, if there exists no proper monotone extension of the graph of A on $\mathcal{H} \times \mathcal{H}$. Let us mention that in case A is maximally monotone, $\operatorname{zer} A$ is a convex and closed set [4, Proposition 23.39]. We refer to [4, Section 23.4] for conditions ensuring that $\operatorname{zer} A$ is nonempty.

The operator A is said to be γ-*strongly monotone* with $\gamma > 0$, if $\langle x - y, u - v \rangle \geq \gamma\|x - y\|^2$ for all $(x, u), (y, v) \in \operatorname{Gr} A$. Notice that if A is maximally monotone and strongly monotone, then $\operatorname{zer} A$ is a singleton, thus nonempty (see [4, Corollary 23.37]).

Let $\gamma > 0$ be arbitrary. A single-valued operator $A : \mathcal{H} \to \mathcal{H}$ is said to be γ-*cocoercive*, if $\langle x-y, Ax-Ay\rangle \geq \gamma\|Ax-Ay\|^2$ for all $(x,y) \in \mathcal{H}\times\mathcal{H}$, and γ-*Lipschitz continuous*, if $\|Ax-Ay\| \leq \gamma\|x-y\|$ for all $(x,y) \in \mathcal{H}\times\mathcal{H}$.

The *resolvent* of A, $J_A : \mathcal{H} \rightrightarrows \mathcal{H}$, is defined by $J_A = (\operatorname{Id} + A)^{-1}$, where $\operatorname{Id} : \mathcal{H} \to \mathcal{H}, \operatorname{Id}(x) = x$ for all $x \in \mathcal{H}$, is the *identity operator* on $\mathcal{H}$. Moreover, if A is maximally monotone, then $J_A : \mathcal{H} \to \mathcal{H}$ is single-valued and maximally monotone (cf. [4, Proposition 23.7 and Corollary 23.10]). For an arbitrary $\gamma > 0$ we have (see [4, Proposition 23.18])

$$J_{\gamma A} + \gamma J_{\gamma^{-1}A^{-1}} \circ \gamma^{-1} \operatorname{Id} = \operatorname{Id}. \tag{5}$$

When $f \in \Gamma(\mathcal{H})$ and $\gamma > 0$, for every $x \in \mathcal{H}$ we denote by $\operatorname{prox}_{\gamma f}(x)$ the *proximal point* of parameter γ of f at x, which is the unique optimal solution of the optimization problem

$$\inf_{y\in\mathcal{H}} \left\{ f(y) + \frac{1}{2\gamma}\|y-x\|^2 \right\}. \tag{6}$$

Notice that $J_{\gamma\partial f} = (\operatorname{Id}_{\mathcal{H}} + \gamma\partial f)^{-1} = \operatorname{prox}_{\gamma f}$, thus $\operatorname{prox}_{\gamma f} : \mathcal{H} \to \mathcal{H}$ is a single-valued operator fulfilling the extended *Moreau's decomposition formula*

$$\operatorname{prox}_{\gamma f} + \gamma \operatorname{prox}_{(1/\gamma)f^*} \circ \gamma^{-1} \operatorname{Id}_{\mathcal{H}} = \operatorname{Id}_{\mathcal{H}}. \tag{7}$$

Let us also recall that the function $f : \mathcal{H} \to \overline{\mathbb{R}}$ is said to be γ-*strongly convex* for $\gamma > 0$, if $f - \frac{\gamma}{2}\|\cdot\|^2$ is a convex function. Let us mention that this property implies γ-strong monotonicity of ∂f (see [4, Example 22.3]).

The *Fitzpatrick function* associated to a monotone operator A, defined as

$$\varphi_A : \mathcal{H}\times\mathcal{H} \to \overline{\mathbb{R}},\ \varphi_A(x,u) = \sup_{(y,v)\in \operatorname{Gr} A} \{\langle x,v\rangle + \langle y,u\rangle - \langle y,v\rangle\},$$

is a convex and lower semicontinuous function and it will play an important role throughout the paper. Introduced by Fitzpatrick in [16], this notion opened the gate towards the employment of convex analysis specific tools when investigating the maximality of monotone operators (see [4–9, 12, 22] and the references therein). In case A is maximally monotone, φ_A is proper and it fulfills

$$\varphi_A(x,u) \geq \langle x,u\rangle\ \forall (x,u) \in \mathcal{H}\times\mathcal{H},$$

with equality if and only if $(x,u) \in \operatorname{Gr} A$. Notice that if $f \in \Gamma(\mathcal{H})$, then ∂f is a maximally monotone operator (cf. [19]) and it holds $(\partial f)^{-1} = \partial f^*$. Furthermore, the following inequality is true (see [5])

$$\varphi_{\partial f}(x,u) \leq f(x) + f^*(u)\ \forall (x,u) \in \mathcal{H}\times\mathcal{H}. \tag{8}$$

We refer the reader to [5], for formulae of the corresponding Fitzpatrick functions computed for particular classes of monotone operators.

We close the section by presenting an algorithm and the corresponding convergence statement for the following monotone inclusion problem, that will be used later in the paper.

Problem 1 Let $\mathcal{H}$ be a real Hilbert space, $A : \mathcal{H} \rightrightarrows \mathcal{H}$ a maximally monotone operator, $D : \mathcal{H} \to \mathcal{H}$ a monotone and η^{-1}-Lipschitz continuous operator with $\eta > 0$, $B : \mathcal{H} \to \mathcal{H}$ a monotone and μ^{-1}-Lipschitz continuous operator with $\mu > 0$ and suppose that $M = \operatorname{zer} B \neq \emptyset$. The monotone inclusion problem to solve is

$$0 \in Ax + Dx + N_M(x).$$

The following forward-backward-forward algorithm for solving Problem 1 was proposed in [10].

Algorithm 2

$$\begin{array}{ll} \texttt{Initialization:} & \texttt{Choose } x_1 \in \mathcal{H} \\ \texttt{For } n \in \mathbb{N} \texttt{ set:} & p_n = J_{\lambda_n A}(x_n - \lambda_n D x_n - \lambda_n \beta_n B x_n) \\ & x_{n+1} = \lambda_n \beta_n (Bx_n - Bp_n) + \lambda_n (Dx_n - Dp_n) + p_n, \end{array}$$

where $(\lambda_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ are sequences of positive real numbers.

For the convergence of the algorithm one needs the following hypotheses:

$$(H_{fitz}) \left\{ \begin{array}{l} (i)\ A + N_M \text{ is maximally monotone and } \operatorname{zer}(A + D + N_M) \neq \emptyset; \\ (ii)\ \text{ For every } p \in \operatorname{ran} N_M, \sum_{n\in\mathbb{N}} \lambda_n \beta_n \left[\sup\limits_{u \in M} \varphi_B\left(u, \frac{p}{\beta_n}\right) - \sigma_M\left(\frac{p}{\beta_n}\right) \right] < +\infty; \\ (iii)\ (\lambda_n)_{n\in\mathbb{N}} \in \ell^2 \setminus \ell^1. \end{array} \right.$$

Remark 3 The first part of the statement in (i) is verified if one of the Rockafellar conditions $M \cap \operatorname{int} \operatorname{dom} A \neq \emptyset$ or $\operatorname{dom} A \cap \operatorname{int} M \neq \emptyset$, is fulfilled (see [20]). We refer the reader to [4, 6–9, 22] for further conditions which guarantee the maximality of the sum of maximally monotone operators. Further, we refer to [4, Subsection 23.4] for conditions eunsuring that the set of zeros of a maximally monotone operator is nonempty. According to [10, Remark 5], the hypothesis (ii) is a generalization of the condition considered in [3] (see also (H_{fitz}^{opt}) and Remark 11 in Section 3 for conditions guaranteeing (ii)).

Before stating the convergence result, we need the following notation. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and $(\lambda_k)_{k\in\mathbb{N}}$ a sequence of positive numbers such that $\sum_{k\in\mathbb{N}} \lambda_k = +\infty$. Let $(z_n)_{n\in\mathbb{N}}$ be the sequence of weighted averages defined as (see [3])

$$z_n = \frac{1}{\tau_n} \sum_{k=1}^{n} \lambda_k x_k, \text{ where } \tau_n = \sum_{k=1}^{n} \lambda_k \ \forall n \in \mathbb{N}. \tag{9}$$

The proof of the following theorem was given in [10] and relies on Fejér-type monotonicity techniques.

Theorem 4 *(see [10, Theorems 20 and 21]) Let $(x_n)_{n\in\mathbb{N}}$ and $(p_n)_{n\in\mathbb{N}}$ be the sequences generated by Algorithm 2 and $(z_n)_{n\in\mathbb{N}}$ the sequence defined in* (9)*. If (H_{fitz}) is fulfilled and $\limsup_{n\to+\infty}\left(\frac{\lambda_n\beta_n}{\mu} + \frac{\lambda_n}{\eta}\right) < 1$, then $(z_n)_{n\in\mathbb{N}}$ converges weakly to an element in $\operatorname{zer}(A + D + N_M)$ as $n \to +\infty$. If, additionally, A is strongly monotone, then $(x_n)_{n\in\mathbb{N}}$ converges strongly to the unique element in $\operatorname{zer}(A + D + N_M)$ as $n \to +\infty$.*

2 Tseng's type penalty schemes

In this section we propose a forward-backward-forward algorithm for solving the following involving linearly composed and parallel-sum type monotone operators and investigate its convergence.

Problem 5 Let $\mathcal{H}$ be a real Hilbert space, $A : \mathcal{H} \rightrightarrows \mathcal{H}$ a maximally monotone operator and $C : \mathcal{H} \to \mathcal{H}$ a monotone and ν-Lipschitz continuous operator for $\nu > 0$. Let m be a strictly positive integer and for any $i \in \{1, ..., m\}$ let $\mathcal{G}_i$ be a real Hilbert space, $B_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ a maximally monotone operator, $D_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ a monotone operator such that D_i^{-1} is ν_i-Lipschtz continuous for $\nu_i > 0$ and $L_i : \mathcal{H} \to \mathcal{G}_i$ a nonzero linear continuous operator. Consider also $B : \mathcal{H} \to \mathcal{H}$ a monotone and μ^{-1}-Lipschitz continuous operator with $\mu > 0$ and suppose that $M = \operatorname{zer} B \neq \emptyset$. The monotone inclusion problem to solve is

$$0 \in Ax + \sum_{i=1}^{m} L_i^*(B_i \Box D_i)(L_i x) + Cx + N_M(x). \tag{10}$$

Let us present our algorithm for solving this problem.

Algorithm 6

$$\begin{array}{ll}
\texttt{Initialization}: & \text{Choose } (x_1, v_{1,1}, ..., v_{m,1}) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m \\
\texttt{For } n \in \mathbb{N} \texttt{ set}: & p_n = J_{\lambda_n A}[x_n - \lambda_n(Cx_n + \sum_{i=1}^m L_i^* v_{i,n}) - \lambda_n \beta_n B x_n] \\
& q_{i,n} = J_{\lambda_n B_i^{-1}}[v_{i,n} + \lambda_n(L_i x_n - D_i^{-1} v_{i,n})],\ i = 1, ..., m \\
& x_{n+1} = \lambda_n \beta_n (Bx_n - Bp_n) + \lambda_n(Cx_n - Cp_n) \\
& \qquad\qquad + \lambda_n \sum_{i=1}^m L_i^*(v_{i,n} - q_{i,n}) + p_n \\
& v_{i,n+1} = \lambda_n L_i(p_n - x_n) + \lambda_n(D_i^{-1} v_{i,n} - D_i^{-1} q_{i,n}) + q_{i,n}, i = 1, ..., m,
\end{array}$$

where $(\lambda_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ are sequences of positive real numbers.

Remark 7 In case $Bx = 0$ for all $x \in \mathcal{H}$, Algorithm 6 collapses into the error-free variant of the iterative scheme given in [14, Theorem 3.1] for solving the monotone inclusion problem

$$0 \in Ax + \sum_{i=1}^{m} L_i^*(B_i \Box D_i)(L_i x) + Cx,$$

since in this case $M = \mathcal{H}$, hence $N_M(x) = \{0\}$ for all $x \in \mathcal{H}$.

For the convergence result we need the following additionally hypotheses (we refer the reader to the remarks 3 and 11 for sufficient conditions guaranteeing $(H_{fitz}^{par-sum})$):

$$(H_{fitz}^{par-sum}) \left\{ \begin{array}{l}
(i)\ A + N_M \text{ is maximally monotone and} \\
\operatorname{zer}\left(A + \sum_{i=1}^m L_i^* \circ (B_i \Box D_i) \circ L_i + C + N_M\right) \neq \emptyset; \\
(ii) \text{ For every } p \in \operatorname{ran} N_M, \sum_{n\in\mathbb{N}} \lambda_n \beta_n \left[\sup_{u\in M} \varphi_B\left(u, \frac{p}{\beta_n}\right) - \sigma_M\left(\frac{p}{\beta_n}\right)\right] < +\infty; \\
(iii)\ (\lambda_n)_{n\in\mathbb{N}} \in \ell^2 \setminus \ell^1.
\end{array} \right.$$

Let us give the main statement of this section. The proof relies on the fact that Problem 5 can be written in the same form as Problem 1 in an appropriate product space.

Theorem 8 *Consider the sequences generated by Algorithm 6 and* $(z_n)_{n\in\mathbb{N}}$ *the sequence defined in* (9). *If* $(H_{fitz}^{par-sum})$ *is fulfilled and* $\limsup_{n\to+\infty}\left(\frac{\lambda_n\beta_n}{\mu}+\lambda_n\beta\right)<1$, *where*

$$\beta=\max\{\nu,\nu_1,...,\nu_m\}+\sqrt{\sum_{i=1}^m\|L_i\|^2},$$

then $(z_n)_{n\in\mathbb{N}}$ *converges weakly to an element in* $\operatorname{zer}\left(A+\sum_{i=1}^m L_i^*\circ(B_i\Box D_i)\circ L_i+C+N_M\right)$ *as* $n\to+\infty$. *If, additionally,* A *and* B_i^{-1}, $i=1,...,m$ *are strongly monotone, then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly to the unique element in* $\operatorname{zer}\left(A+\sum_{i=1}^m L_i^*\circ(B_i\Box D_i)\circ L_i+C+N_M\right)$ *as* $n\to+\infty$.

Proof. We start the proof by noticing that $x\in\mathcal{H}$ is a solution to Problem 5 if and only if there exist $v_1\in\mathcal{G}_1,...,v_m\in\mathcal{G}_m$ such that

$$\left\{\begin{array}{l} 0\in Ax+\sum_{i=1}^m L_i^*v_i+Cx+N_M(x)\\ v_i\in(B_i\Box D_i)(L_ix), i=1,...,m,\end{array}\right. \tag{11}$$

which is nothing else than

$$\left\{\begin{array}{l} 0\in Ax+\sum_{i=1}^m L_i^*v_i+Cx+N_M(x)\\ 0\in B_i^{-1}v_i+D_i^{-1}v_i-L_ix, i=1,...,m.\end{array}\right. \tag{12}$$

In the following we endow the product space $\mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m$ with inner product and associated norm defined for all $(x,v_1,...,v_m),(y,w_1,...,w_m)\in\mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m$ as

$$\langle(x,v_1,...,v_m),(y,w_1,...,w_m)\rangle=\langle x,y\rangle+\sum_{i=1}^m\langle v_i,w_i\rangle$$

and

$$\|(x,v_1,...,v_m)\|=\sqrt{\|x\|^2+\sum_{i=1}^m\|v_i\|^2},$$

respectively.

We introduce the operators $\widetilde{A}:\mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m\rightrightarrows\mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m$

$$\widetilde{A}(x,v_1,...,v_m)=Ax\times B_1^{-1}v_1\times....\times B_m^{-1}v_m,$$

$\widetilde{D}:\mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m\to\mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m$,

$$\widetilde{D}(x,v_1,...,v_m)=\Big(\sum_{i=1}^m L_i^*v_i+Cx,D_1^{-1}v_1-L_1x,...,D_m^{-1}v_m-L_mx\Big)$$

and $\widetilde{B}:\mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m\to\mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m$,

$$\widetilde{B}(x,v_1,...,v_m)=(Bx,0,...,0).$$

Notice that, since A and B_i, $i=1,...,m$ are maximally monotone, $\widetilde{A}$ is maximally monotone, too (see [4, Props. 20.22, 20.23]). Further, as it was done in [14, Theorem 3.1], one can show that $\widetilde{D}$ is a monotone and β-Lipschitz continuous operator. For the sake of completeness we include here some details of the proof of these two statements.

Let be $(x, v_1, ..., v_m), (y, w_1, ..., w_m) \in \mathcal{H}\times \mathcal{G}_1 \times ... \times \mathcal{G}_m$. By using the monotonicity of C and D_i^{-1}, $i = 1, ..., m$. we have

$$\begin{aligned}&\langle (x, v_1, ..., v_m) - (y, w_1, ..., w_m), \widetilde{D}(x, v_1, ..., v_m) - \widetilde{D}(y, w_1, ..., w_m)\rangle \\ &\qquad = \langle x - y, Cx - Cy\rangle + \sum_{i=1}^{m} \langle v_i - w_i, D_i^{-1} v_i - D_i^{-1} w_i\rangle \\ &\qquad + \sum_{i=1}^{m} (\langle x - y, L_i^*(v_i - w_i)\rangle - \langle v_i - w_i, L_i(x - y)\rangle) \geq 0,\end{aligned}$$

which shows that $\widetilde{D}$ is monotone.

The Lipschitz continuity of $\widetilde{D}$ follows by noticing that

$$\begin{aligned}&\left\| \widetilde{D}(x, v_1, ..., v_m) - \widetilde{D}(y, w_1, ..., w_m)\right\| \\ &\leq \left\| \left(Cx - Cy, D_1^{-1} v_1 - D_1^{-1} w_1, ..., D_m^{-1} v_m - D_m^{-1} w_m\right)\right\| \\ &+ \left\| \left(\sum_{i=1}^{m} L_i^*(v_i - w_i), -L_1(x - y), ..., -L_m(x - y)\right)\right\| \\ &\leq \sqrt{\nu^2 \|x - y\|^2 + \sum_{i=1}^{m} \nu_i^2 \|v_i - w_i\|^2} + \sqrt{\left(\sum_{i=1}^{m} \|L_i\| \cdot \|v_i - w_i\|\right)^2 + \sum_{i=1}^{m} \|L_i\|^2 \cdot \|x - y\|^2} \\ &\leq \beta \|(x, v_1, ..., v_m) - (y, w_1, ..., w_m)\|.\end{aligned}$$

Moreover, $\widetilde{B}$ is monotone, μ^{-1}-Lipschitz continuous and

$$\operatorname{zer} \widetilde{B} = \operatorname{zer} B \times \mathcal{G}_1 \times ... \times \mathcal{G}_m = M \times \mathcal{G}_1 \times ... \times \mathcal{G}_m,$$

hence

$$N_{\widetilde{M}}(x, v_1, ..., v_m) = N_M(x) \times \{0\} \times ... \times \{0\},$$

where

$$\widetilde{M} = M \times \mathcal{G}_1 \times ... \times \mathcal{G}_m = \operatorname{zer} \widetilde{B}.$$

Taking into consideration (12), we obtain that $x \in \mathcal{H}$ is a solution to Problem 5 if and only if there exist $v_1 \in \mathcal{G}_1, ..., v_m \in \mathcal{G}_m$ such that

$$(x, v_1, ..., v_m) \in \operatorname{zer}(\widetilde{A} + \widetilde{D} + N_{\widetilde{M}}).$$

Conversely, when $(x, v_1, ..., v_m) \in \operatorname{zer}(\widetilde{A} + \widetilde{D} + N_{\widetilde{M}})$, then $x \in \operatorname{zer}\left(A + \sum_{i=1}^{m} L_i^* \circ (B_i \Box D_i) \circ L_i + C + N_M\right)$. This means that determining the zeros of $\widetilde{A} + \widetilde{D} + N_{\widetilde{M}}$ will automatically provide a solution to Problem 5.

Using that

$$J_{\lambda\widetilde{A}}(x, v_1, ..., v_m) = (J_{\lambda A_1}(x), J_{\lambda B_1^{-1}}(v_1), ..., J_{\lambda B_m^{-1}}(v_m))$$

for every $(x, v_1, ..., v_m) \in \mathcal{H}\times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ and every $\lambda > 0$ (see [4, Proposition 23.16]), one can easily see that the iterations of Algorithm 6 read for any $n \in \mathbb{N}$:

$$\left\{ \begin{array}{l} (p_n, q_{1,n}, ..., q_{m,n}) = J_{\lambda_n \widetilde{A}} \Big[(x_n, v_{1,n}, ..., v_{m,n}) - \lambda_n \widetilde{D}(x_n, v_{1,n}, ..., v_{m,n}) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad -\lambda_n\beta_n \widetilde{B}(x_n, v_{1,n}, ..., v_{m,n})\Big] \\ (x_{n+1}, v_{1,n+1}, ..., v_{m,n+1}) = \lambda_n\beta_n \Big[\widetilde{B}(x_n, v_{1,n}, ..., v_{m,n}) - \widetilde{B}(p_n, q_{1,n}, ..., q_{m,n})\Big] \\ \qquad\qquad +\lambda_n \Big[\widetilde{D}(x_n, v_{1,n}, ..., v_{m,n}) - \widetilde{D}(p_n, q_{1,n}, ..., q_{m,n})\Big] + (p_n, q_{1,n}, ..., q_{m,n}), \end{array} \right.$$

which is nothing else than the iterative scheme of Algorithm 2 employed to the monotone inclusion problem

$$0 \in \widetilde{A}x + \widetilde{D}x + N_{\widetilde{M}}(x).$$

In order to compute the Fitzpatrick function of $\widetilde{B}$, we consider arbitrary elements $(x, v_1, ..., v_m), (x', v'_1, ..., v'_m) \in \mathcal{H}\times \mathcal{G}_1 \times ... \times \mathcal{G}_m$. It holds

$$\begin{aligned} &\varphi_{\widetilde{B}}\left((x, v_1, ..., v_m), (x', v'_1, ..., v'_m)\right) = \\ &\sup_{\substack{(y,w_1,...,w_m)\in \\ \mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m}} \Big\{\langle(x, v_1, ..., v_m), \widetilde{B}(y, w_1, ..., w_m)\rangle + \langle(x', v'_1, ..., v'_m), (y, w_1, ..., w_m)\rangle \\ &\qquad\qquad - \langle(y, w_1, ..., w_m), \widetilde{B}(y, w_1, ..., w_m)\rangle\Big\} \\ &= \sup_{\substack{(y,w_1,...,w_m)\in \\ \mathcal{H}\times\mathcal{G}_1\times...\times\mathcal{G}_m}} \left\{\langle x, By\rangle + \langle x', y\rangle + \sum_{i=1}^m \langle v'_i, w_i\rangle - \langle y, By\rangle\right\}, \end{aligned}$$

thus

$$\varphi_{\widetilde{B}}\big((x, v_1, ..., v_m), (x', v'_1, ..., v'_m)\big) = \left\{ \begin{array}{ll} \varphi_B(x, x'), & \text{if } v'_1 = ... = v'_m = 0, \\ +\infty, & \text{otherwise.} \end{array} \right.$$

Moreover,

$$\sigma_{\widetilde{M}}(x, v_1, ..., v_m) = \left\{ \begin{array}{ll} \sigma_M(x), & \text{if } v_1 = ... = v_m = 0, \\ +\infty, & \text{otherwise,} \end{array} \right.$$

hence condition (ii) in $(H^{par-sum}_{fitz})$ is nothing else than

for each $(p, p_1, ..., p_m) \in \operatorname{ran} N_{\widetilde{M}} = \operatorname{ran} N_M \times \{0\} \times ... \times \{0\}$

$$\sum_{n\in\mathbb{N}} \lambda_n\beta_n \left[\sup_{(u,v_1,...,v_m)\in\widetilde{M}} \varphi_{\widetilde{B}}\left((u, v_1, ..., v_m), \frac{(p, p_1, ..., p_m)}{\beta_n}\right) - \sigma_{\widetilde{M}}\left(\frac{(p, p_1, ..., p_m)}{\beta_n}\right)\right] < +\infty.$$

Moreover, condition (i) in $(H^{par-sum}_{fitz})$ ensures that $\widetilde{A} + N_{\widetilde{M}}$ is maximally monotone and $\operatorname{zer}(\widetilde{A}+\widetilde{D}+N_{\widetilde{M}}) \neq \emptyset$. Hence, we are in the position of applying Theorem 4 in the context of finding the zeros of $\widetilde{A}+\widetilde{D}+N_{\widetilde{M}}$. The statements of the theorem are an easy consequence of this result. ■

3 Convex minimization problems

In this section we employ the results given for monotone inclusion when minimizing a convex function with an intricate formulation with respect to the set of minimizers of a convex and differentiable function with Lipschitz continuous gradient.

Problem 9 Let $\mathcal{H}$ be a real Hilbert space, $f \in \Gamma(\mathcal{H})$ and $h : \mathcal{H} \to \mathbb{R}$ a convex and differentiable function with a ν-Lipschitz continuous gradient for $\nu > 0$. Let m be a strictly positive integer and for any $i = 1, ..., m$ let $\mathcal{G}_i$ be a real Hilbert space, $g_i, l_i \in \Gamma(\mathcal{G}_i)$ such that l_i is ν_i^{-1}-strongly convex for $\nu_i > 0$ and $L_i : \mathcal{H} \to \mathcal{G}_i$ a nonzero linear continuous operator. Further, let $\Psi \in \Gamma(\mathcal{H})$ be differentiable with a μ^{-1}-Lipschitz continuous gradient, fulfilling $\min \Psi = 0$. The convex minimization problem under investigation is

$$\inf_{x \in \operatorname{argmin} \Psi} \left\{ f(x) + \sum_{i=1}^{m} (g_i \Box l_i)(L_i x) + h(x) \right\}. \tag{13}$$

Consider the maximal monotone operators

$$A = \partial f, B = \nabla \Psi, C = \nabla h, B_i = \partial g_i \text{ and } D_i = \partial l_i, i = 1, ..., m.$$

According to [4, Proposition 17.10, Theorem 18.15], $D_i^{-1} = \nabla l_i^*$ is a monotone and ν_i-Lipschitz continuous operator for $i = 1, ..., m$. Moreover, B is a monotone and μ^{-1}-Lipschitz continuous operator and

$$M := \operatorname{argmin} \Psi = \operatorname{zer} B.$$

Taking into account the sum rules of the convex subdifferential, every element of $\operatorname{zer}\left(\partial f + \sum_{i=1}^{m} L_i^* \circ (\partial g_i \Box \partial l_i) \circ L_i + \nabla h + N_M\right)$ is an optimal solution of (13). The converse is true if an appropriate qualification condition is satisfied. For the readers convenience, let us present some qualification conditions which are suitable in this context. One of the weakest qualification conditions of interiority-type reads (see, for instance, [14, Proposition 4.3, Remark 4.4])

$$(0, ..., 0) \in \operatorname{sqri}\left(\prod_{i=1}^{m} \operatorname{dom} g_i - \{(L_1 x, ..., L_m x) : x \in \operatorname{dom} f \cap M\}\right). \tag{14}$$

The condition (14) is fulfilled if (i) $\operatorname{dom} g_i = \mathcal{G}_i$, $i = 1, ..., m$ or (ii) $\mathcal{H}$ and $\mathcal{G}_i$ are finite-dimensional and there exists $x \in \operatorname{ri} \operatorname{dom} f \cap \operatorname{ri} M$ such that $L_i x \in \operatorname{ri} \operatorname{dom} g_i$, $i = 1, ..., m$ (see [14, Proposition 4.3]).

Algorithm 6 becomes in this particular case

Algorithm 10

`Initialization:` `Choose` $(x_1, v_{1,1}, ..., v_{m,1}) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$

`For` $n \in \mathbb{N}$ `set:`

$$\begin{aligned}
p_n &= \operatorname{prox}_{\lambda_n f}[x_n - \lambda_n(\nabla h(x_n) + \textstyle\sum_{i=1}^{m} L_i^* v_{i,n}) - \lambda_n \beta_n \nabla \Psi(x_n)] \\
q_{i,n} &= \operatorname{prox}_{\lambda_n g_i^*}[v_{i,n} + \lambda_n(L_i x_n - \nabla l_i^*(v_{i,n}))],\ i = 1, ..., m \\
x_{n+1} &= \lambda_n \beta_n (\nabla \Psi(x_n) - \nabla \Psi(p_n)) + \lambda_n (\nabla h(x_n) - \nabla h(p_n)) \\
&\quad + \lambda_n \textstyle\sum_{i=1}^{m} L_i^*(v_{i,n} - q_{i,n}) + p_n \\
v_{i,n+1} &= \lambda_n L_i(p_n - x_n) + \lambda_n(\nabla l_i^*(v_{i,n}) - \nabla l_i^*(q_{i,n})) + q_{i,n},\ i = 1, ..., m
\end{aligned}$$

For the convergence result we need the following hypotheses:

$$(H_{fitz}^{opt}) \begin{cases} (i)\ \partial f + N_M \text{ is maximally monotone and (13) has an optimal solution;} \\ (ii)\ \text{For every } p \in \operatorname{ran} N_M, \sum_{n \in \mathbb{N}} \lambda_n \beta_n \left[\Psi^*\left(\frac{p}{\beta_n}\right) - \sigma_M\left(\frac{p}{\beta_n}\right)\right] < +\infty; \\ (iii)\ (\lambda_n)_{n \in \mathbb{N}} \in \ell^2 \setminus \ell^1. \end{cases}$$

Remark 11 (a) Let us mention that $\partial f + N_M$ is maximally monotone, if $0 \in \operatorname{sqri}(\operatorname{dom} f - M)$, a condition which is fulfilled if, for instance, f is continuous at a point in $\operatorname{dom} f \cap M$ or $\operatorname{int} M \cap \operatorname{dom} f \neq \emptyset$.

(b) Since $\Psi(x) = 0$ for all $x \in M$, by (8) it follows that whenever (ii) in (H_{fitz}^{opt}) holds, condition (ii) in $(H_{fitz}^{par-sum})$, formulated for $B = \nabla\Psi$, is also true.

(c) Let us mention that hypothesis (ii) is satisfied, if $\sum_{n\in\mathbb{N}} \frac{\lambda_n}{\beta_n} < +\infty$ and Ψ is bounded below by a multiple of the square of the distance to C (see [2]). This is for instance the case when $M = \operatorname{zer} L = \{x \in \mathcal{H} : Lx = 0\}$, $L : \mathcal{H} \to \mathcal{H}$ is a linear continuous operator with closed range and $\Psi : \mathcal{H} \to \mathbb{R}, \Psi(x) = \|Lx\|^2$ (see [2, 3]). For further situations for which condition (ii) is fulfilled we refer to [3, Section 4.1].

We are able now to formulate the convergence result.

Theorem 12 *Consider the sequences generated by Algorithm 10 and $(z_n)_{n\in\mathbb{N}}$ the sequence defined in* (9)*. If (H_{fitz}^{opt}) and* (14) *are fulfilled and $\limsup_{n\to+\infty}\left(\frac{\lambda_n\beta_n}{\mu} + \lambda_n\beta\right) < 1$, where*

$$\beta = \max\{\nu, \nu_1, ..., \nu_m\} + \sqrt{\sum_{i=1}^{m} \|L_i\|^2},$$

then $(z_n)_{n\in\mathbb{N}}$ converges weakly to an optimal solution to (13) *as $n \to +\infty$. If, additionally, f and g_i^*, $i = 1, ..., m$ are strongly convex, then $(x_n)_{n\in\mathbb{N}}$ converges strongly to the unique optimal solution of* (13) *as $n \to +\infty$.*

Remark 13 (a) According to [4, Proposition 17.10, Theorem 18.15], for a function $g \in \Gamma(\mathcal{H})$ one has that g is strongly convex if and only if g is differentiable with Lipschitz continuous gradient.

(b) Notice that in case $\Psi(x) = 0$ for all $x \in \mathcal{H}$, Algorithm 10 turns out to be the error-free variant of the iterative scheme given in [14, Theorem 4.2] for solving the convex minimization problem

$$\inf_{x\in\mathcal{H}} \left\{ f(x) + \sum_{i=1}^{m} (g_i \Box l_i)(L_i x) + h(x) \right\}. \tag{15}$$

4 A numerical experiment in TV-based image inpainting

In this section we illustrate the applicability of Algorithm 10 when solving an image inpainting problem, which aims for recovering lost information. We consider images of size $M \times N$ as vectors $x \in \mathbb{R}^n$ for $n = M \cdot N$, while each pixel denoted by $x_{i,j}$, $1 \leq i \leq M$, $1 \leq j \leq N$, ranges in the closed interval from 0 (pure black) to 1 (pure white). We denote by $b \in \mathbb{R}^n$ the image with missing pixels (in our case set to black) and by $K \in \mathbb{R}^{n\times n}$ the diagonal matrix with $K_{i,i} = 0$, if the pixel i in the noisy image $b \in \mathbb{R}^n$ is missing, and $K_{i,i} = 1$, otherwise, $i = 1, ..., n$ (notice that $\|K\| = 1$). The original image will be reconstructed by considering the following TV-regularized model

$$\inf\left\{TV_{\text{iso}}(x) : Kx = b, x \in [0, 1]^n\right\}. \tag{16}$$

The objective function $TV_{\text{iso}} : \mathbb{R}^n \to \mathbb{R}$ is the isotropic total variation defined by

$$TV_{\text{iso}}(x) = \sum_{i=1}^{M-1}\sum_{j=1}^{N-1} \sqrt{(x_{i+1,j}-x_{i,j})^2 + (x_{i,j+1}-x_{i,j})^2}$$
$$+ \sum_{i=1}^{M-1} |x_{i+1,N} - x_{i,N}| + \sum_{j=1}^{N-1} |x_{M,j+1} - x_{M,j}| .$$

We show that (16) can be written as an optimization problem of type (9). To this end, we denote $\mathcal{Y} = \mathbb{R}^n \times \mathbb{R}^n$ and define the linear operator $L : \mathbb{R}^n \to \mathcal{Y}$, $x_{i,j} \mapsto (L_1 x_{i,j}, L_2 x_{i,j})$, where

$$L_1 x_{i,j} = \begin{cases} x_{i+1,j} - x_{i,j}, & \text{if } i < M \\ 0, & \text{if } i = M \end{cases} \text{ and } L_2 x_{i,j} = \begin{cases} x_{i,j+1} - x_{i,j}, & \text{if } j < N \\ 0, & \text{if } j = N \end{cases} .$$

The operator L represents a discretization of the gradient in the horizontal and vertical directions. One can easily check that $\|L\|^2 \leq 8$ and that its adjoint $L^* : \mathcal{Y} \to \mathbb{R}^m$ is as easy to implement as the operator itself (cf. [13]). Further, for $(y,z),(p,q) \in \mathcal{Y}$, one can introduce the inner product on $\mathcal{Y}$

$$\langle (y,z),(p,q)\rangle = \sum_{i=1}^{M}\sum_{j=1}^{N}(y_{i,j}p_{i,j} + z_{i,j}q_{i,j}),$$

which induces a norm defined as $\|(y,z)\|_\times = \sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{y_{i,j}^2 + z_{i,j}^2}$. One can see that $TV_{\text{iso}}(x) = \|Lx\|_\times$ for every $x \in \mathbb{R}^n$.

Further, by considering the function $\Psi : \mathbb{R}^n \to \mathbb{R}$, $\Psi(x) = \frac{1}{2}\|Kx - b\|^2$, problem (16) can be reformulated as

$$\inf_{x \in \operatorname{argmin} \Psi} \{f(x) + g_1(Lx)\}, \tag{17}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $f = \delta_{[0,1]^n}$ and $g_1 : \mathcal{Y} \to \mathbb{R}$, $g_1(y_1, y_2) = \|(y_1, y_2)\|_\times$. Problem (17) is of type (9), when one takes $m = 1$, $L_1 = L$, $l_1 = \delta_{\{0\}}$ and $h = 0$. Notice that $\nabla\Psi(x) = K(Kx - b) = K(x - b)$ for every $x \in \mathbb{R}^n$, thus $\nabla\Psi$ is Lipschitz continuous with Lipschitz constant $\mu = 1$. The iterative scheme in Algorithm 10 becomes for every $n \geq 0$ in this case

$$\begin{cases} p_n = \operatorname{prox}_{\lambda_n f}[x_n - \lambda_n L^* v_{1,n} - \lambda_n \beta_n K(x_n - b)]; \\ q_{1,n} = \operatorname{prox}_{\lambda_n g_1^*}(v_{1,n} + \lambda_n L x_n); \\ x_{n+1} = \lambda_n \beta_n K(x_n - p_n) + \lambda_n L^*(v_{1,n} - q_{1,n}) + p_n; \\ v_{1,n+1} = \lambda_n L(p_n - x_n) + q_{1,n}. \end{cases}$$

For the proximal points we have the following formulae:

$$\operatorname{prox}_{\gamma f}(x) = \operatorname{proj}_{[0,1]^n}(x) \ \forall \gamma > 0 \text{ and } \forall x \in \mathbb{R}^n$$

and (see [11])

$$\operatorname{prox}_{\gamma g_1^*}(p,q) = \operatorname{proj}_S(p,q) \ \ \forall \gamma > 0 \text{ and } \forall (p,q) \in \mathcal{Y},$$

where

$$S = \left\{ (p,q) \in \mathcal{Y} : \max_{\substack{1 \le i \le M \\ 1 \le j \le N}} \sqrt{p_{i,j}^2 + q_{i,j}^2} \le 1 \right\}$$

and the projection operator $\mathrm{proj}_S : \mathcal{Y} \to S$ is defined via

$$(p_{i,j}, q_{i,j}) \mapsto \frac{(p_{i,j}, q_{i,j})}{\max\left\{1, \sqrt{p_{i,j}^2 + q_{i,j}^2}\right\}}, \ 1 \le i \le M, \ 1 \le j \le N.$$

We tested the algorithm on the fruit image and considered as parameters $\lambda_n = 0.9 \cdot n^{-0.75}$ and $\beta_n = n^{0.75}$ for any $n \in \mathbb{N}$. Figure 1 shows the original image, the image obtained from it after setting 80% randomly chosen pixels to black, the nonaveraged reconstructed image x^n and the averaged reconstructed image z^n after 1000 iterations.

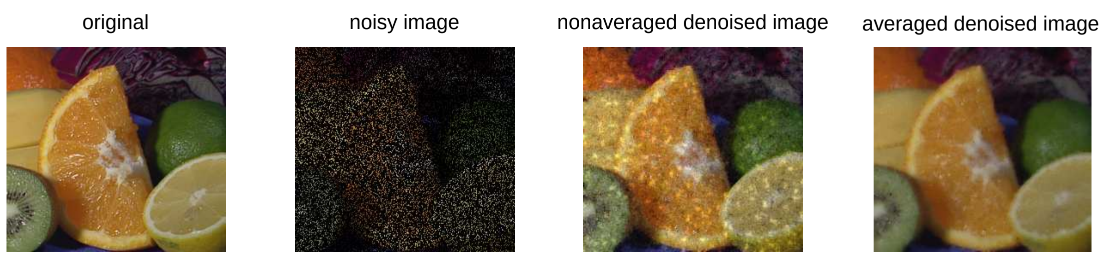

Figure 1: TV image inpainting: the original image, the image with 80% missing pixels, the nonaveraged reconstructed image x^n and the averaged reconstructed image z^n after 1000 iterations

The comparisons concerning the quality of the reconstructed images were made by means of the improvement in signal-to-noise ratio (ISNR), which is defined as

$$\mathrm{ISNR}(n) = 10 \log_{10} \left(\frac{\|x - b\|^2}{\|x - x^n\|^2} \right),$$

where x, b and x^n denote the original, the image with missing pixels and the recovered image at iteration n, respectively.

Figure 2 shows the evolution of the ISNR values for the averaged and the nonaveraged reconstructed images. Both figures illustrate the theoretical outcomes concerning the sequences involved in Theorem 12, namely that the averaged sequence has better convergence properties than the nonaveraged one.

Acknowledgements. The authors are thankful to the anonymous reviewers for their recommendations which improved the quality of the paper.

References

[1] H. Attouch, M.-O. Czarnecki, *Asymptotic behavior of coupled dynamical systems with multiscale aspects*, Journal of Differential Equations 248(6), 1315-1344, 2010

[2] H. Attouch, M.-O. Czarnecki, J. Peypouquet, *Prox-penalization and splitting methods for constrained variational problems*, SIAM Journal on Optimization 21(1), 149-173, 2011

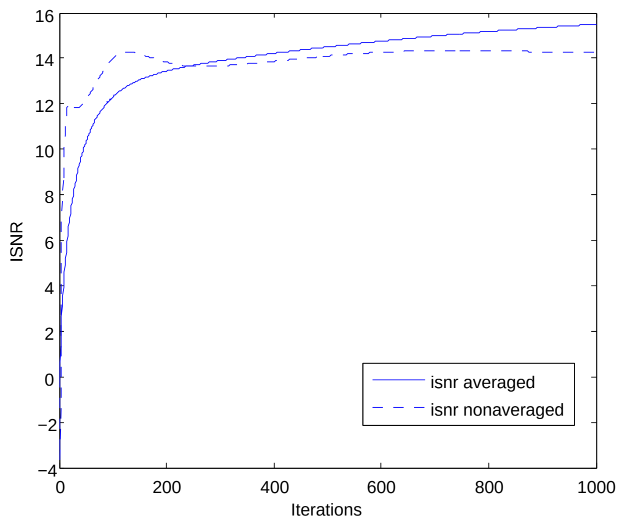

Figure 2: The ISNR curves for the averaged and nonaveraged reconstructed images

[3] H. Attouch, M.-O. Czarnecki, J. Peypouquet, *Coupling forward-backward with penalty schemes and parallel splitting for constrained variational inequalities*, SIAM Journal on Optimization 21(4), 1251-1274, 2011

[4] H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York, 2011

[5] H.H. Bauschke, D.A. McLaren, H.S. Sendov, *Fitzpatrick functions: inequalities, examples and remarks on a problem by S. Fitzpatrick*, Journal of Convex Analysis 13(3-4), 499–523, 2006

[6] J.M. Borwein, *Maximal monotonicity via convex analysis*, Journal of Convex Analysis 13(3-4), 561–586, 2006

[7] J.M. Borwein and J.D. Vanderwerff, *Convex Functions: Constructions, Characterizations and Counterexamples*, Cambridge University Press, Cambridge, 2010

[8] R.I. Boţ, *Conjugate Duality in Convex Optimization*, Lecture Notes in Economics and Mathematical Systems, Vol. 637, Springer, Berlin Heidelberg, 2010

[9] R.I. Boţ, E.R. Csetnek, *An application of the bivariate inf-convolution formula to enlargements of monotone operators*, Set-Valued Analysis 16(7-8), 983–997, 2008

[10] R.I. Boţ, E.R. Csetnek, *Forward-backward and Tseng's type penalty schemes for monotone inclusion problems*, arXiv:1306.0352, 2013

[11] R.I. Boţ, C. Hendrich, *Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization*, arXiv:1211.1706v1 [math.OC], 2012

[12] R.S. Burachik, B.F. Svaiter, *Maximal monotone operators, convex functions and a special family of enlargements*, Set-Valued Analysis 10(4), 297–316, 2002

[13] A. Chambolle, *An algorithm for total variation minimization and applications*, Journal of Mathematical Imaging and Vision, 20(1–2), 89–97, 2004

[14] P.L. Combettes, J.-C. Pesquet, *Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators*, Set-Valued and Variational Analysis 20(2), 307–330, 2012

[15] I. Ekeland, R. Temam, *Convex Analysis and Variational Problems*, North-Holland Publishing Company, Amsterdam, 1976

[16] S. Fitzpatrick, *Representing monotone operators by convex functions*, in: Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988), Proceedings of the Centre for Mathematical Analysis **20**, Australian National University, Canberra, 59–65, 1988

[17] N. Noun, J. Peypouquet, *Forward-backward penalty scheme for constrained convex minimization without inf-compactness*, Journal of Optimization Theory and Applications, 158(3), 787–795, 2013

[18] J. Peypouquet, *Coupling the gradient method with a general exterior penalization scheme for convex minimization*, Journal of Optimization Theory and Applications 153(1), 123-138, 2012

[19] R.T. Rockafellar, *On the maximal monotonicity of subdifferential mappings*, Pacific Journal of Mathematics 33(1), 209–216, 1970

[20] R.T. Rockafellar, *On the maximality of sums of nonlinear monotone operators*, Transactions of the American Mathematical Society 149, 75–88, 1970

[21] R.T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM Journal on Control and Optimization 14(5), 877-898, 1976

[22] S. Simons, *From Hahn-Banach to Monotonicity*, Springer, Berlin, 2008

[23] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002